Annals of Epidemiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Commentary

Considerations for transgender population health research based on US national surveys

Elle Lett, PhD, MA[a,b,c,*], Avery Everhart, MA[a,d]

[a] Center for Applied Transgender Studies, Chicago, IL
[b] Center for Health Equity Advancement, University of Pennsylvania, Philadelphia, PA
[c] Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
[d] Dornsife Spatial Sciences Institute, University of Southern California, Los Angeles, CA

ARTICLE INFO

Article history:
Received 29 June 2021
Revised 24 September 2021
Accepted 22 October 2021
Available online xxx

Keywords:
Transgender
Minority health
Surveys and questionnaires
Bias
Health inequities
Population health
Demography

ABSTRACT

Transgender identities and health are highly politicized in the United States leading to restrictions on relevant data collection in national health surveillance systems. This has serious implications on transgender population health research; an urgent area of study given the systemic discrimination faced by transgender individuals and the resultant social and health inequities. In this precarious political climate, obtaining high-quality data for research is challenging and in recent years, two data sources have formed the foundation of transgender health research in the United States, namely the 2015 United States Transgender Study and the Behavioral Risk Factor Surveillance System after the launch of the optional Sexual Orientation and Gender Identity Module in 2014. While useful, there are serious challenges to using these data to study transgender health, specifically related to survey weighting methodologies, ascertainment of gender identity, and study design. In this article, we detail these challenges and discuss the strengths and weaknesses of various methodological approaches that have been implemented as well as clarify several common errors that exist in the literature. We feel that this contribution is necessary to provide accurate interpretation of the evidence that currently informs policy and priority setting for transgender population health and will provide vital insights for future studies with these now ubiquitous sources of data in the field.

© 2021 Elsevier Inc. All rights reserved.

Transgender individuals, or individuals whose gender identity differs from their assigned gender at birth (AGAB) [1], face stigma and discrimination that drive social and health inequities in this population relative to cisgender persons in the United States [2]. Social inequities include employment discrimination, income inequality, homelessness, victimization and incarceration [3], and health inequities include depression, anxiety, suicidality, post-traumatic stress disorder, and substance use disorders, among others, all relative to the cisgender US population [4–9]. These findings have been replicated across many study designs and data sources that generally fall into the following groups (1) single hospital or multi-institution health system studies using primary survey data collection or secondary analysis of electronic health records (EHR) [7,10,11], (2) secondary analysis of large administrative/insurance claims databases [12,13], (3) analysis of national surveys with piloted modules for identifying transgender identity [14–17], and (4) the 2015 United States Transgender Survey [3]. The last two sources, which ostensibly have the potential to provide the most regionally diverse estimates of the health status in the transgender population, will be the focus of this article. We will emphasize challenges to using these data sources based on their (1) ascertainment of transgender identity and (2) weighting procedure and methods to address them, as well as provide insights into interpreting past and designing future studies that use these valuable data.

Data sources

One of the challenges to studying transgender population health in the United States is that transgender identity is not ascertained in the Census or most national health surveys. This has

Abbreviations: BRFSS, behavioral risk factor surveillance system; USTS, united states transgender survey; SOGI, Sexual Orientation and Gender Identity; ACS, American Community Survey; NH, Native Hawaiian; PI, Pacific Islander.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

* Corresponding author. E. Lett, Center for Health Equity Advancement, University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104.
E-mail address: ellel@upenn.edu (E. Lett).

https://doi.org/10.1016/j.annepidem.2021.10.009
1047-2797/© 2021 Elsevier Inc. All rights reserved.

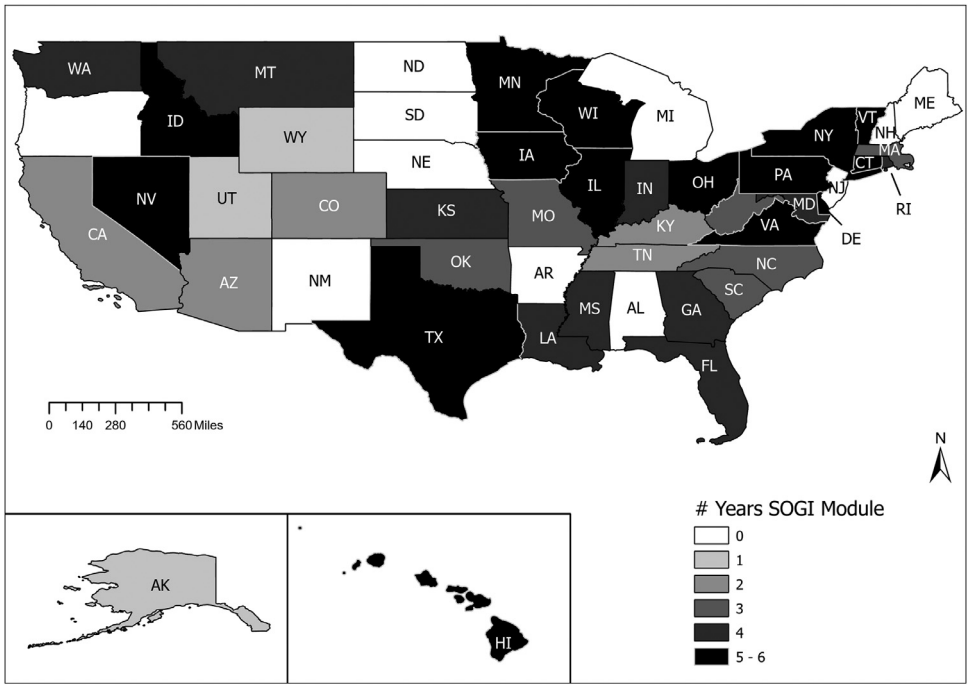

Fig. 1.. Map of US states depicting frequency of BRFSS SOGI module administration, 2014–2019.

Table 1
Number of states administering the SOGI module by year, 2014–2019

2014	2015	2016	2017	2018	2019
19	22	25	27	29	30

been linked to the excessive politicization of transgender health care in the United States, with efforts to include transgender individuals in national surveys advancing or contracting based on the political ideologies of the executive administration [18,19]. The most notable exception to this exclusion is the Behavioral Risk Factor Surveillance System (BRFSS), one of the longest running annual health survey in the United States. The BRFSS is a telephone-based survey that uses a complex survey design that uses stratified cluster sampling, and poststratification weight adjustment with raking to approximate the overall US population. During 2014, BRFSS piloted an optional sexual orientation and gender identity (SOGI) module that includes a question about transgender identity. Between 2014 and 2019, a combined total of 38 states have administered this module at least once (Table 1), though only 6 have administered it in all years and 12 states have never used the SOGI module. Importantly, this means that while the overall BRFSS datasets are indeed nationally representative, the subsamples of transgender respondents are inadequate for generalizing to national patterns (Fig. 1).

Other notable surveys that include transgender identity include the Youth Risk Behavior Surveillance System (since 2017, selected districts) [17], National Crime Victimization Survey (since 2016) [20], the General Social Survey (since 2018) [21], and National HIV Behavioral Surveillance System (since 2019, selected cities) [22].

In contrast to BRFSS, the other major, national data source used for transgender population health studies is the 2015 US Transgender Survey. This was a web-based survey administered by the National Center for Transgender Equality, that had 27,715 respondents, making it the largest study of the US transgender population in history. Unlike BRFSS, this survey only includes transgender individuals, and, as a convenience sample, was not designed to represent the entire US transgender population. Further, the USTS has respondents from every state and territory, but is not necessarily nationally representative in the sense of other surveys like BRFSS which employs a complex survey design to sample individuals from prespecified clusters within strata to accurately reflect the overall national population.

Ascertaining transgender identity

There have been several studies with the goal of developing survey items that reliably identify transgender individuals. One difficulty of asking this question relates to the diversity of transgender identities a well as the acceptability of asking questions related to gender identity. For the former, transgender people include those who identify with an individual gender identity, such as transgender women and transgender men, as well as those who identify along or outside of the gender spectrum, such as nonbinary, nonconforming, and agender individuals. There is added complexity that some people may not identify with the term transgender despite living as a gender that is different than their AGAB. For health surveillance purposes these individuals are still among the target transgender population but are not necessarily captured in asking the same questions.

The two-question format for capturing gender identity date back to 1997, as noted by Reisner et al. [23], but were first tested for efficacy against a one-question format by Charlotte Chuck Tate et al. “A Two-Question Method for Assessing Gender Category in Social and Medical Sciences” [24]. Their studies compared two formats, shown in Box 1.

Box 1. Tate et al. one-question and two-question gender identity survey item formats

One-Question Format:
What is your gender? Male, Female, Transgender, Other
Two-Question Format:

What is your current gender identity?: Female, Male, Transgender, Genderqueer
What gender where you assigned at birth?: Female, Male, Intersex

In their study, Tate et al. evaluated the formats in terms of missing data, predictive validity, and percent of survey respondents identified as transgender, across two studies of university students and one study of community members. The predictive validity was assessed by comparing self-reported gender identity in a prior in-person interview, with the gender identity obtained from the instruments. They found that the one-question format used in the anonymous survey had higher missing data (1.26% among university students) than the two-question format (0% among university students and 0.16% among community members) and lower predictive validity. Further, the two-question format identified nearly two times as many transgender individuals (1.65% among university students, and 3.09% among community members compared to 0.84% for the one-question format among university students) across different samples in university and community settings.

Since the 2013 Tate study, several studies have tested multiple versions of two-question format in different contexts, assessing the acceptability of asking the question among cisgender and transgender populations, and different ethnoracial groups [25–28]. Because of its higher predictive performance and capture rate of transgender individuals, the two-question format has received greater attention among subject matter experts and therefore has evolved to be more inclusive of more contemporary gender terms including gender nonconforming, and gender nonbinary individuals. Despite the evidence of improved response rate and higher predictive validity with greater yield of identifying transgender individuals, the BRFSS, YRBS, and all the CDC national surveys listed above still use a one-step method (see below) [29]. This introduces misclassification error for the transgender population that may bias estimates when comparing transgender to cisgender groups. In the case of binary outcomes this will generally bias estimates toward the null, but in the context of polytomous outcomes, or in the presence of effect modification, interaction or mediation, the direction of bias is not always predictable [30]. Further, to quantify the severity of any bias that may occur due to misclassification, it would be necessary to evaluate the predictive performance of both question formats among a cohort of individuals with variation in a health outcome of interest. To our knowledge, there are no publicly available datasets that contain sufficient information to facilitate such a bias quantification, so future studies are necessary to further characterize the performance of these transgender identity ascertainment methods for health surveillance. Additionally, large-scale studies comparing the feasibility and acceptability of the formats across transgender and cisgender populations on a national level have yet to be reported, which are important to facilitate broad uptake of these gender identity questions across different survey systems. Despite these open questions, based on the increased capture of transgender individuals—a relatively rare but increasing population in the United States [31], the current scholarly consensus is to use the two-step format (Box 2) [1].

Box 2. BRFSS transgender survey item, 2014–2019

2. Do you consider yourself to be transgender?
If yes, ask "Do you consider yourself to be (1) Male-to-female, (2) Female-to-male, or (3) Gender nonconforming?

Interviewer note: Please say the number before the yes text response. Respondent can answer with either the number of the text/word.
1 Yes, Transgender, male-to-female
2 Yes, Transgender, female to male
3 Yes, Transgender, gender nonconforming
4 No
7 Don't know/not sure
9 Refused

The USTS survey only includes transgender individuals so the issues of misclassification in the BRFSS are not relevant to for that data source. However, it cannot be used to directly estimate comparisons between transgender and cisgender groups so selecting between these two data sources requires a careful consideration of the study question and balancing the need for comparison with the potential of bias due to misclassification. USTS has the added benefit of being designed specifically for transgender individuals such that the instrument includes more relevant questions that may not be available in BRFSS.

Weighting procedures

Behavioral risk factor surveillance system survey weights

As the largest probabilistic sample of transgender individuals, the BRFSS has been an attractive source for transgender health studies and has been used extensively to study mental health, behavioral risk factors, self-rated health, and chronic conditions in this population [16,32,33]. For these studies, the authors typically default to using design-weighted analyses based on the weights provided with the BRFSS data. These weights are calculated using a poststratification adjustment algorithm known as raking or iterative proportional fitting, whereby the weights are re-estimated to fit a target marginal distribution of selected demographics, including sex [34]. However, sex in the BRFSS, up to and including 2015, was not asked from the respondent but rather assigned by the interviewer based on their assessment of the vocal timbre of the interviewee. Two studies have demonstrated that this is subject to significant misclassification among transgender individuals [35,36]. Box 3 contains the item related to sex included in BRFSS from 2015.

Box 3. BRFSS sex item (2015)

8.21 Indicate sex of respondent. Ask only if necessary.
1 Male [*Go to* Q8.23]
2 Female [*If respondent is* 45 *years old or older, go to* Q8.23].

Depending on how this question is interpreted, this misclassifies assigned gender at birth by 30% [36] or current gender identity by 70% [35]. Because the raking algorithm incorporates sex, this essentially embeds misclassification into the weights for transgender individuals. It should also be noted that it is possible that this approach of identifying the sex may also misclassify cisgender individuals, but the degree to which this occurs is not recoverable. It is only observable because transgender respondents can self-identity as "male-to-female" or "female-to-male" in BRFSS, allowing for assigned gender at birth and current gender to be identified. (It's important to note that this language is considered outdated and stigmatizing but was useful in the context of addressing this question).

In subsequent years since 2015, the BRFSS questionnaire has contained ambiguous instructions for interviewers on how to codify sex which have included embedding presumed sex into early-stage questions before respondents are given the opportunity to self-identify [29]. Further, if an interviewee refuses to give their sex when prompted, the interviewers are instructed to terminate the call. From 2016 to 2018 interviewers were instructed to code for sex implicitly in selecting eligibility responses in both the cell phone and landline instructions, specifically when asking if the interviewee is 18 years of age or older [29]. The questionnaire did not instruct interviewers to ask respondents to self-identify their sex until the demographics section, which was the eighth of seventeen core sections. Most modules in the questionnaire that have skips built in based on sex only appear after the demographics section; however, a notable exception is the question about diabetes in section 6 where only those respondents that the interviewer assumes to be female and who answer yes to having ever been diagnosed with diabetes are then asked if it was only during pregnancy. The questionnaire changed in 2019 to include an eligibility question that explicitly instructs the interviewer to ask the interviewee if they are male or female [29].

The 2015 United States Transgender Survey eights

In contrast, the USTS survey samples transgender people from all states in the US and in that sense can be considered nationally inclusive. However, one serious limitation of this data is its poor ethnoracial diversity, which is cited in many studies using the USTS [37–40]. Below is a table that shows selected demographics of the USTS sample as well as weights they generated to "adjust the USTS sample to reflect the race/ethnicity, age, and educational attainment of the US population." It is important to draw a distinction between the survey weights included with the BRFSS and the survey weights included with the USTS. The former are weights that are based on the probability a respondent is sampled under the complex cluster stratified design of the BRFSS. Application of these weights rely on design-based inference (also known as randomization inference) which assumes a random sample from a finite population in which the probability of sampling each observation is precisely known [34]. Analyses based on this inference can be easily implemented with specialized packages within most statistical software such as the *survey* [41] package in R, or the survey [42] commands in Stata. In contrast, the USTS weights are a *post hoc* adjustment that assumes a simple random sample (SRS). These weights can be used in the default functions for fitting generalized linear models. Unfortunately, there may be some confusion about this distinction with at least one example of an investigator inappropriately using methods relying on design-based inference with the USTS weights, which would lead to invalid standard error estimates [43].

As shown above, the ethnoracial distribution of the sample is very homogenous with the majority of the sample, 82.53%, being white, with severe underrepresentation of Black/African American and Hispanic or Latino populations relative to the US. The USTS generated separate weights for each of the above demographics and took the product of these weights as the final survey weight for each observation (separately for education above and below 25). The main limitation of this approach is that it assumes that a reasonable target distribution is the general US population, and that the USTS sample is a random sample of the transgender population. For the former, there is evidence that the distribution of ethnoracial identities may differ between transgender and cisgender groups [16,44], that educational attainment is impeded in transgender groups[44], likely due in part to discrimination [45], and that the transgender population is increasing in younger age groups [46]. For the latter assumption, given the socioeconomic inequities experienced by the transgender population and the fact that it is a web-based survey, there is potentially a substantial selection bias that makes the assumptions underlying these weights dubious.

Methodological alternatives

Approaches to analyzing BRFSS data for transgender populations

To date, three approaches have been implemented for analysis focused on transgender health in the BRFSS. (1) The most common approach is to conduct design-weighted analysis using the survey weights provided with the data [16,33]. (2) Another approach is to exclude the survey weights and only account for strata and clustering in the sampling schema of the survey [47]. (3) The last approach has been applied for comparisons of transgender groups to cisgender groups and involves using matching algorithms, such as exact matching or propensity score matching, to develop matched sets of transgender and cisgender respondents for answering the study questions [14,15,49]. These approaches have key differences in terms of how they contend with the misclassification error and potential bias introduced by the raked BRFSS weights, and the target estimand of the study.

The first approach, design-weighted analyses, ignores the misclassification error and assumes that the weights are unbiased. Depending on the study design and framing, with gender identity as a proxy, the target estimand is the average association between exposure to systemic transphobia and the study outcome in the United States. Given the limitations of the BRFSS, this target estimand is not strictly attainable; only 38 states have fielded the SOGI module, and have done so for different time windows, introducing potential temporal (i.e., sampling the transgender population of respective states at different times) and sampling biases (i.e., sampling the transgender population of only a subset of US states). One could make the case that the target estimand is instead the average association only in the subset of states that have fielded the SOGI module, but the utility of such an estimand for national population health surveillance and policy efforts is questionable, and the temporal bias still holds. Therefore, it is unclear what benefits are gained in exchange for the misclassification error introduced by the survey weights.

The second approach, excluding the raked weights and adjusting for stratified and clustered sampling only, will generally, for the same study design, have similar target estimands as the first approach. However, in contrast to the first approach which uses the design-based inference to directly attain national level estimates using the weights, this approach essentially treats the BRFSS as an unweighted cluster stratified random sample of the United States [34]. The analysis still relies on design-based inference but rather than reweighting the data to obtain national level estimates, it is left to the investigator and reader to determine how generalizable the findings are to the broader US transgender population. By avoiding the raked weights, any bias due to the misclassification error introduced from sex-specific poststratification adjustment is avoided.

The last approach, matching transgender individuals to cisgender individuals based on selected respondent characteristics, ignores the survey design completely and treats the BRFSS as a random sample. The main implication of this approach is that generally, it will have a different target estimand depending on the matching approach. One commonly used approach, propensity score matching transgender individuals to sets of cisgender individuals, would generally estimate the "average exposure effect on the exposed"(AEE), rather than the overall effect of transgender status on the outcome of interest or the "average exposure effect in the population" [50,51]. There are approaches with and without

Table 2
Ethnoracial identity, age, and educational attainment of 2015 United States transgender survey respondents and calculated survey weights

Sociodemographic characteristic	USTS (%)	2014 ACS* (%)	Weight ($\frac{ACS\%}{USTS\%}$)
Education, 18–24			
Less than high school	5.66	13.90	2.46
High school graduate	19.32	30.20	1.56
Some college or associate's degree	58.97	45.80	0.78
Bachelor's degree or higher	16.06	10.10	0.63
Education, 25 and older			
Less than high school	1.49	13.10	8.81
High school graduate	7.43	27.70	3.73
Some college	25.50	21.00	0.82
Associate's degree	11.27	8.20	0.73
Bachelor's degree	32.45	18.70	0.58
Graduate or professional degree	21.87	11.40	0.52
Age			
18–24 y	42.72	12.83	0.30
25–44 y	39.64	34.26	0.86
45–64 y	14.74	34.06	2.31
65 y and older	2.90	18.85	6.51
Ethnoracial Group			
Native American	1.15	0.66	0.58
Asian/NH/PI alone	2.84	5.10	1.80
Two or more races/some other race alone	5.30	2.33	0.44
Black or African American alone	2.87	12.24	4.26
Hispanic or Latino alone	5.31	16.90	3.18
White alone	82.53	62.77	0.76

NH = Native Hawaiian; PI = Pacific Islander.
Demographics data provided in the 2015 USTS codebook per data use agreement between author (AE) and the Interuniversity Consortium on Political and Social Research with approval from IRB at the University of Southern California.
High school graduate includes individuals who passed General Education Development (GED) tests.
Native American was originally referred to as American Indian or Alaska Native in the USTS documentation.
* Only among ACS respondents 18 years old and older.

propensity scores that target other estimands outside of the AEE including the average exposure effect in the population such as full matching [51], but these have not been applied to transgender population health studies using the BRFSS [14,15,49]. In the context of transgender population health, under certain assumptions, the AEE can be interpreted as the causal effect of transgender status and the corresponding discrimination on health outcomes, estimated only among transgender individuals. As with the second approach, matching abandons the goal of obtaining estimates representative of all states included in the subsample that fielded the SOGI module and avoids the misclassification error that is introduced by the raked weights. Additionally, unless sampling variables are included in the matching algorithm, clustering is also ignored and this approach essentially reduces the BRFSS to a large, regionally diverse, convenience sample of transgender and cisgender individuals that are useful for comparisons and identifying gender-based heath inequities.

Approaches to analyzing 2015 United States transgender survey

With the USTS data, unweighted approaches and [38,52] weighted [37,39,40,53,54], are used to analyze the USTS data. The weighted approaches are more common and are applied to unadjusted descriptive statistics, such as summary tables describing the demographic information (the traditional "Table 1" in epidemiologic studies), as well as regression analyses to estimate exposure-outcome relationships after covariate adjustment. Confusingly, there is even a study that presented unweighted counts and weighted percentages in their demographics tables which is misleading [54]; the authors are essentially presenting partial information for two different samples. It would be more appropriate to present the complete unweighted data to give an accurate representation of the actual observed sample; and, when weighted analyses are justified, present the new distribution of the artificial sample that is created by applying the weights. However, to our knowledge this is done in a relatively few studies using the USTS data [40].

The main problem with the common practice of defaulting to presenting only the weighted analysis of the USTS data is that it obscures the critical flaw of the 2015 USTS data; namely, its poor representation of the ethnoracial diversity of the transgender population. Overall the transgender population faces significant health inequities, but when one applies an intersectional lens [55–57], it becomes clear that ethnoracial minoritized subpopulations are subject to different and sometimes greater challenges [14,15]. Further, Black and Hispanic transgender individuals face economic and health inequities that are often more severe than other ethnoracial groups [14,15], making barriers to web-based surveys such as the USTS more likely, and selection bias more severe amongst these multiply marginalized groups. The implication of using only the weighted data is that we overestimate the precision of estimates within these substrata. As shown in Table 2, the USTS sample contained 2.87% Black respondents, which corresponds to approximately 800 of the near 28,000 transgender people in the study. However, using the weighted data, the sample is treated as if there were approximately 3200 Black respondents. In this way, by relying solely on weighted analyses investigators place too much confidence, in the form of narrower interval estimates and lower P values, in conclusions drawn about the Black transgender population based on the ethnoracially homogenous USTS data, which is subject to an unknowable, but potentially severe degree of selection bias. Unfortunately, that selection bias can only be mitigated in future samples, perhaps with the next iteration of the United States Transgender Survey, which hopefully will have more purposeful recruitment through organizations and online community spaces that cater to transgender people of color. Until then, when using the 2015 USTS data we should not overestimate our certainty in this biased sample by defaulting to the weighted analyses, particularly when conducting analyses focused on poorly represented subpopulations within the data. Rather, we should present

both weighted and unweighted analyses and present the potential sources of bias as they relate to the study question so that the reader can interpret the findings in the appropriate context.

Other data sources

In addition to the previously discussed surveys, administrative claims databases and EHR data have become increasingly common as data sources for the study of transgender population health. Generally, studies with these data sources are not nationally representative; however, for administrative claims data from large payor systems such as MarketScan [12,58] the cohorts may be regionally diverse, and for large networks with integrated EHR systems the cohorts may be representative for a given geographic area depending on the patient population served [11]. Taken together, these studies could potentially advance transgender health surveillance in many settings. However, one key limitation with these alternatives is that most algorithms used for ascertainment of transgender identity in EHR or claims data are based on diagnosis codes for gender dysphoria (formerly known as gender identity disorder) [12,13,59], a disorder characterized by stress related to difference in their gender identity their assigned gender at birth [60], which only affects a subset of transgender individuals. This introduces misclassification error which leads to many of the same problems discussed above with BRFSS. Additionally, there is substantial variation in the implementation of these algorithms, and the authors of this commentary are currently engaged in a systematic review to characterize the rationale and validation of the related approaches [61]. Recently, natural language processing methods of free-text data have been shown to greatly improve the identification of transgender individuals in some EHR systems, but these data are not routinely available in administrative claims data [62,63]. Therefore, until we have universal uptake of new sexual orientation and gender identity data modules mandated for inclusion in electronic health records and similar information is reported in claims databases [19], the promise of high-quality surveillance data for transgender health with these data sources will likely remain unfulfilled.

Conclusion

The goal of this piece was to discuss some of the challenges and methodological approaches to transgender population health studies using the two most dominant large-scale, survey data sources, namely the BRFSS and the 2015 USTS. The BRFSS uses suboptimal identification methods for both gender identity and assigned gender at birth creating misclassification error that may bias at multiple levels from design of the survey instrument to data collection, and from analysis to interpretation of results. In comparison, the USTS is tailored to the transgender population but is subject to selection bias that makes conclusions about ethnoracial minoritized groups within the transgender population hard to generalize from this survey. Overall, the field of transgender population health will benefit greatly from universal SOGI data inclusion of national surveys including the Census, as well as careful consideration of how to obtain more representative samples in convenience web-based studies. However, there are lessons to be learned from other nationally representative surveys which use probabilistic sampling, such as the General Social Survey [21] in the US or even the census enumeration process in countries like Nepal [64], about the pitfalls of implementation without proper training. Still, these data provide useful conclusions when interpreted in context of these limitations and will continue to help motivate important policy advances for transgender health.

Acknowledgments

Elle Lett would like to thank the Black Health Scholars Network for their support of this work. The authors would like to thank Dr. Ayden Scheim for his thoughtful review of an early version of this manuscript. The authors have no funding to report for this work.

References

[1] Kronk CA, Everhart AR, Ashley F, Thompson HM, Schall TE, Goetz TG, et al. Transgender data collection in the electronic health record: current concepts and issues. J Am Med Informatics Assoc 2021;0(June):1–14. doi:10.1093/jamia/ocab136.
[2] White Hughto JM, Reisner SL, Pachankis JE. Transgender stigma and health: a critical review of stigma determinants, mechanisms, and interventions. Soc Sci Med 2015;147:222–31. doi:10.1016/j.socscimed.2015.11.010.
[3] James SE, Herman JL, Rankin S, Keisling M, Mottet L, Anafi M. The report of the 2015 US transgender survey. National Center for Transgender Equality 2016:1–302. Published online https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf .
[4] Reisner SL, Biello KB, Hughto JMW, Kuhns L, Mayer KH, Garofalo R, et al. Psychiatric diagnoses and comorbidities in a diverse, multicity cohort of young transgender women: baseline findings from project LifeSkills. JAMA Pediatr 2016;170(5):481–6. doi:10.1001/jamapediatrics.2016.0067.
[5] Crissman HP, Stroumsa D, Kobernik EK, Berger MB. Gender and frequent mental distress: comparing transgender and non-transgender individuals' self-rated mental health. J Women's Heal 2019;28(2):143–51. doi:10.1089/jwh.2018.7411.
[6] Lefevor GT, Boyd-Rogers CC, Sprague BM, Janis RA. Health disparities between genderqueer, transgender, and cisgender individuals: an extension of minority stress theory. J Couns Psychol 2019;66(4):385–95. doi:10.1037/cou0000339.
[7] Brown GR, Jones KT. Mental health and medical health disparities in 5135 transgender veterans receiving healthcare in the Veterans Health Administration: a case-control study. LGBT Heal 2016;3(2):122–31. doi:10.1089/lgbt.2015.0058.
[8] Su D, Irwin JA, Fisher C, et al. Mental health disparities within the LGBT population: a comparison between transgender and nontransgender individuals. Transgender Heal 2016;1(1):12–20. doi:10.1089/trgh.2015.0001.
[9] Witcomb GL, Pierre W, Claes L, Brewin N, Crawford JR, Arcelus J. Levels of depression in transgender people and its predictors: results of a large matched control study with transgender people accessing clinical services. J Affect Disord 2018;235:308–15. doi:10.1016/j.jad.2018.02.051.
[10] Kauth MR, Shipherd JC, Lindsay J, Blosnich JR. Access to care for transgender veterans in the Veterans Health Administration : 2006 –2013. Am J Public Health 2014;104:2013–15. doi:10.2105/AJPH.2014.302086.
[11] Getahun D, Nash R, Flanders WD, et al. Cross-sex hormones and acute cardiovascular events in transgender persons: a cohort study. Ann Intern Med 2018;169(4):205–13. doi:10.7326/M17-2785.
[12] McDowell A, Progovac AM, Cook BL, Rose S. Estimating the health status of privately insured gender minority children and adults. LGBT Heal 2019;6(6):289–96. doi:10.1089/lgbt.2018.0238.
[13] Proctor K, Haffer SC, Ewald E, Hodge C, James CV. Identifying the transgender population in the Medicare program. Transgender Heal 2016;1(1):250–65. doi:10.1089/trgh.2016.0031.
[14] Lett E, Dowshen NL, Baker KE. Intersectionality and health inequities for gender minority blacks in the U.S. Am J Prev Med 2020;59(5):639–47. doi:10.1016/j.amepre.2020.04.013.
[15] Lett E, Asabor EN, Beltrán S, Dowshen N. Characterizing health inequities for the U.S. transgender hispanic population using the behavioral risk factor surveillance system. Transgender Heal 2020;00(00):1–9. doi:10.1089/trgh.2020.0095.
[16] Baker KE. Findings from the behavioral risk factor surveillance system on health-related quality of life among US transgender adults, 2014-2017. JAMA Intern Med 2019;179(8):1141–4. doi:10.1001/jamainternmed.2018.7931.
[17] Johns MM, Lowry R, Andrzejewski J, et al. Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students — 19 states and large urban school districts, 2017. Morb Mortal Wkly Rep 2019;68(3):67–71. doi:10.15585/MMWR.MM6803A3.
[18] Cahill SR, Makadon HJ. If they don't count us, we don't count: Trump administration rolls back sexual orientation and gender identity data collection. LGBT Heal 2017;4(3):171–3. doi:10.1089/lgbt.2017.0073.
[19] Baker KE, Streed CG, Durso LE. Ensuring that LGBTQI+ people count — collecting data on sexual orientation, gender identity, and intersex status. N Engl J Med 2021 Published online. doi:10.1056/NEJMp2032447.
[20] Flores AR, Meyer IH, Langton L, Herman JL. Gender identity disparities in criminal victimization: national crime victimization survey, 2017-2018. Am J Public Health 2021;111(4):726–9. doi:10.2105/AJPH.2020.306099.
[21] Lagos D, Compton D. Evaluating the use of a two-step gender identity measure in the 2018 general social survey. Demography 2021;58(2):763–72. doi:10.1215/00703370-8976151.
[22] Centers for Disease Control and Prevention. HIV infection, risk, prevention, and testing behaviors among transgender women-national HIV behavioral surveillance, 7 U.S. cities, 2019-2020; 2021. https://www.cdc.gov/hiv/pdf/

library/reports/surveillance/cdc-hiv-surveillance-special-report-number-27.pdf [Accessed 24 August 2021].
[23] Reisner SL, Conron KJ, Scout, et al. Counting" transgender and gender-nonconforming adults in health research: recommendations from the gender identity in US surveillance group. TSQ Transgender Stud Q 2015;2(1):34–57. doi:10.1215/23289252-2848877.
[24] Tate CC, Ledbetter JN, Youssef CP. A two-question method for assessing gender categories in the social and medical sciences. J Sex Res 2013;50(8):767–76. doi:10.1080/00224499.2012.690110.
[25] Cahill S, Singal R, Grasso C, et al. Do ask, do tell: high levels of acceptability by patients of routine collection of sexual orientation and gender identity data in four diverse American community health centers. PLoS ONE 2014;9(9). doi:10.1371/journal.pone.0107104.
[26] Lombardi E, Banik S. The utility of the two-step gender measure within Trans and Cis populations. Sex Res Soc Policy 2016;13(3):288–96. doi:10.1007/s13178-016-0220-6.
[27] Tordoff DM, Morgan J, Dombrowski JC, Golden MR, Barbee LA. Increased ascertainment of transgender and non-binary patients using a 2-step versus 1-step gender identity intake question in an STD clinic setting. Sex Transm Dis 2019;46(4):254–9. doi:10.1097/OLQ.0000000000000952.
[28] Bauer GR, Braimoh J, Scheim AI, Dharma C. Transgender-inclusive measures of sex /gender for population surveys : mixed- methods evaluation and recommendations. PLoS ONE 2017:1–28 Published online.
[29] Centers for Disease Control and Prevention. BRFSS questionaires. Behavioral risk factor surveillance system. Accessed June 28, 2021. https://www.cdc.gov/brfss/questionnaires/index.htm
[30] Van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. Int J Epidemiol 2020;49(1):338–47. doi:10.1093/ije/dyz251.
[31] Zhang Q, Goodman M, Adams N, et al. Epidemiological considerations in transgender health : a systematic review with focus on higher quality data. Int J Transgender Heal 2020;21(2):125–37. doi:10.1080/26895269.2020.1753136.
[32] Streed CG, McCarthy EP, Haas JS. Association between gender minority status and self-reported physical and mental health in the United States. JAMA Intern Med 2017;177(8):1210–12. doi:10.1001/jamainternmed.2017.1460.
[33] Streed CG Jr, McCarthy EP, Haas JS. Self-reported physical and mental health of gender nonconforming transgender adults in the United States. LGBT Heal 2018;5(7):443–8.
[34] Lumley T. Complex surveys: a guide to analysis using R, 565. John Wiley & Sons; 2011.
[35] Cicero EC, Reisner SL, Merwin EI, Humphreys JC, Silva SG. Application of behavioral risk factor surveillance system sampling weights to transgender health measurement. Nurs Res 2020 Published online. doi:10.1097/NNR.0000000000000428.
[36] Tordoff D, Andrasik M, Hajat A. Misclassification of sex assigned at birth in the behavioral risk factor surveillance system and transgender reproductive health: a quantitative bias analysis. Epidemiology 2019;30(5):669–78. doi:10.1097/EDE.0000000000001046.
[37] Kachen A, Pharr JR. Health care access and utilization by transgender populations: a united states transgender survey study. Transgender Heal 2020;5(3):141–8. doi:10.1089/trgh.2020.0017.
[38] Turban JL, Loo SS, Almazan AN, Keuroghlian AS. Factors leading to "Detransition" among transgender and gender diverse people in the united states: a mixed-methods analysis. LGBT Heal 2021;8(4):273–80. doi:10.1089/lgbt.2020.0437.
[39] Becerra MB, Rodriquez EJ, Avina RM, Id JB. Experiences of violence and mental health outcomes among Asian American transgender adults in the United States. PLoS ONE 2021:1–13 Published online. doi:10.1371/journal.pone.0247812.
[40] Kattari SK, Bakko M, Hecht HK, Kattari L. SSM - population health correlations between healthcare provider interactions and mental health among transgender and nonbinary adults. SSM - Popul Heal 2020;10:100525. doi:10.1016/j.ssmph.2019.100525.
[41] Lumley T. Analysis of complex survey samples. J Stat Softw 2004;9:1–19. doi:10.18637/jss.v009.i08.
[42] Kreuter F, Valliant R. A survey on survey statistics : what is done and can be done in Stata Issues in analyzing survey data. State J 2007;7(1):1–21.
[43] Yockey A, King K, Vidourek R, et al. Archives of suicide research past-year suicidal ideation among transgender individuals in the United States past-year suicidal ideation among transgender individuals in the United States. Arch Suicide Res 2020;0(0):1–11. doi:10.1080/13811118.2020.1803165.
[44] Meyer IH, Brown TNT, Herman JL, Reisner SL, Bockting WO. Demographic characteristics and health status of transgender adults in select US regions: behavioral risk factor surveillance system, 2014. Am J Public Health 2017;107(4):582–9. doi:10.2105/AJPH.2016.303648.
[45] Kosciw JG, Greytak EA, Zongrone AD, Clark CM, Truong NL. The 2017 national school climate survey: the experiences of lesbian, gay, bisexual, transgender, and queer youth in our nation's schools. ERIC; 2018.
[46] Zhang Q, Goodman M, Adams N, et al. Epidemiological considerations in transgender health: a systematic review with focus on higher quality data. Int J Transgender Heal 2020;21(2):125–37. doi:10.1080/26895269.2020.1753136.
[47] Cicero EC, Reisner SL, Merwin EI, Humphreys JC, Silva SG. The health status of transgender and gender nonbinary adults in the United States. PLoS ONE 2020;15(2):1–20. doi:10.1371/journal.pone.0228765.
[49] Pharr JR, Batra K. Propensity score analysis assessing the burden of non-communicable diseases among the transgender population in the United States using the behavioral risk factor surveillance system (2017 –2019). Healthcare 2021 Published online.
[50] Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. BMJ 2019;367:1–10. doi:10.1136/bmj.l5657.
[51] Greifer N, Stuart EA. Matching methods for confounder adjustment: an addition to the Epidemiologist's Toolbox. Epidemiol Rev 2021 Published online. doi:10.1093/epirev/mxab003.
[52] Turban JL, King D, Carswell JM, Keuroghlian AS. Pubertal suppression for transgender youth and risk of suicidal ideation. Pediatrics 2020;145(2):1–8.
[53] Kcomt L, Evans-Polce RJ, Boyd CJ, McCabe SE. Association of transphobic discrimination and alcohol misuse among transgender adults: results from the U.S. Transgender Survey. Drug Alcohol Depend 2020;215(July):108223. doi:10.1016/j.drugalcdep.2020.108223.
[54] Stroumsa D, Crissman HP, Dalton VK, Kolenic G, Richardson CR. Insurance coverage and use of hormones among transgender respondents to a national survey. Ann Fam Med 2020;18(6):528–34. doi:10.1370/afm.2586.
[55] Crenshaw K. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. Univ Chic Leg Forum 1989:139–68 Published online.
[56] Crenshaw KW. Mapping the margins: intersectionality, identity politics, and violence against women of color. Stanford Law Rev 1991;43(6):1241–99.
[57] Bowleg L. The problem with the phrase women and minorities: intersectionality-an important theoretical framework for public health. Am J Public Health 2012;102(7):1267–73. doi:10.2105/AJPH.2012.300750.
[58] McDowell A, Raifman J, Progovac AM, Rose S. Association of nondiscrimination policies with mental health among gender minority individuals. JAMA Psychiatry 2020;77(9):952–8. doi:10.1001/jamapsychiatry.2020.0770.
[59] Blosnich JR, Marsiglio MC, Gao S, et al. Mental health of transgender veterans in US States with and without discrimination and hate crime legal protection. Am J Public Health 2016;106(3):534–40. doi:10.2105/AJPH.2015.302981.
[60] Zucker KJ. The DSM-5 diagnostic criteria for gender dysphoria. In: Management of gender dysphoria. Springer; 2015. p. 33–7.
[61] Lett E., Baker K., Kronk C., Streed C. Assessing diagnosis-code based approaches for identifying transgender individuals in electronic health records and administrative claims data: a systematic review. Prospero: international prospective register of systematic reviews. Published 2021. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=238533. [Accessed 30 May 2021].
[62] Xie F, Getahun D, Quinn VP, et al. An automated algorithm using free-text clinical notes to improve identification of transgender people. Informatics Heal Soc Care 2021;46(1):18–28. doi:10.1080/17538157.2020.1828890.
[63] Roblin D, Barzilay J, Tolsma D, et al. A novel method for estimating transgender status using electronic medical records. Ann Epidemiol 2016;26(3):198–203. doi:10.1016/j.annepidem.2016.01.004.
[64] Knight KG, Flores AR, Nezhad SJ. Surveying Nepal's third gender: development, implementation, and analysis. TSQ Transgender Stud Q 2015;2(1):101–22. doi:10.1215/23289252-2848904.